



No evidence for language benefits in infant relational learning

Erin M. Anderson^{a,*}, Yin-Juei Chang^a, Susan Hespos^a, Dedre Gentner^a

^a Department of Psychology, Northwestern University, United States

ARTICLE INFO

Keywords:

Infant cognition
Analogy
Language
Relational learning

ABSTRACT

Recent studies have found that infants show relational learning in the first year. Like older children, they can abstract relations such as *same* or *different* across a series of exemplars. For older children, language has a major impact on relational learning: labeling a shared relation facilitates learning, while labeling component objects can disrupt learning. Here we ask: Does language influence relational learning at 12 months? Experiment 1 ($n = 64$) examined the influence of a relational label on learning. Prior to the study, the infants saw three pairs of objects, all labeled “These are *same*” or “These are *different*”. Experiment 2 ($n = 48$) examined the influence of object labels prior to the study, with three objects labeled (e.g., “This is a cup, this is a tower.”). We compared the present results with those of Ferry et al. (2015), where infants abstracted *same* and *different* relations after undergoing a similar paradigm without prior labels. If the effects of language mirror those in older children, we would expect that infants given relational labels (Experiment 1) will be helped in abstracting *same* and *different* compared to infants not given labels and that infants given object labels (Experiment 2) will be hindered relative to those not given labels. We found no evidence for either prediction. In Experiment 1, infants who had heard relational labels did not benefit compared to infants who had received no labels (Ferry et al., 2015). In Experiment 2, infants who had heard object labels showed the same patterns as those in Ferry et al. (2015), suggesting that object labels had no effect. This finding is important because it highlights a key difference between the relational learning abilities of infants and those seen in older children, pointing to a protracted process by which language and relational learning become entwined.

1. Introduction

Humans are prodigious learners. Not only can we learn by association and by perceptual similarity, but unlike most species, we can learn by relational comparison. That is, humans can detect and transfer relational similarities between situations or ideas, despite perceptual differences. Relational ability has been argued to be the root of higher-order cognition (Gentner, 2003, 2010; Gentner & Medina, 1998; Hofstadter & Sander, 2013). Indeed, there are numerous links between recognizing common relations and higher-order reasoning. Identifying shared structure is critical in categorizing items together or separately, as well as in tracking the hierarchical relationships between categories (e.g., peppers and carrots are both types of vegetable, and vegetables and meat are both types of food). By generalizing relational structure between scenarios, we can also make new discoveries (e.g., that an atom has an orbital structure similar to the solar system). We can even reason about events that have never occurred, by generalizing causal relationships

* Correspondence to: Department of Psychological & Brain Sciences, 1101 E. 10th Street, Bloomington, IN 47405, United States.

E-mail addresses: ema1@iu.edu (E.M. Anderson), alixyj@gmail.com (Y.-J. Chang).

¹ Author’s note: Erin M. Anderson is now a postdoctoral researcher at Indiana University, but conducted this research at Northwestern University.

to hypothetical and counter-factual scenarios (e.g., “If I hadn’t put my cup so close to the edge, it wouldn’t have fallen when I bumped the table.”).

Humans’ relational ability differs dramatically from that of other animals (e.g., Gentner, Shao, Simms & Hespos, 2021; Penn, Holyoak, & Povinelli, 2008). But is this the result of evolutionarily endowed processes, or does it result from our abundant exposure to cultural supports, including language? Even preschool children have already absorbed a great deal of cultural learning. Therefore, recent studies have tested for whether relational ability is present in infancy, probing the roots of this ability before culture and language have a major impact (Addyman & Mareschal, 2010; Anderson, Chang, Hespos, & Gentner, 2018; Ferry, Hespos, & Gentner, 2015; Hochmann et al., 2018; Tyrrell, Stauffer, & Snowman, 1991). Ferry et al. (2015) demonstrated *same-different* learning in 7- and 9-month-olds by habituating infants to pairs that instantiated either *same* or *different* relations. After a learning period of less than ten minutes, infants succeeded in discriminating a novel relation from the relation they had habituated to, even when the test relations included never-before-seen toys. Infants also made a telling failure. Prior to seeing the habituation pairs, infants had been shown a set of individual toys. At test, they failed to discriminate *same* and *different* relations for pairs containing these individually salient toys, despite generalizing to new pairs. Thus, infants showed two key signatures of learning that also characterize relational learning in children and adults. The first is that facilitating comparison across exemplars (here, the habituation pairs) promotes relational abstraction and generalization (Gentner & Christie, 2010; Gentner & Namy, 1999; Gick & Holyoak, 1983). The second is that salient individual objects can disrupt the ability to align and abstract shared relations (Gentner & Rattermann, 1991; Gentner & Toupin, 1986; Markman & Gentner, 1993; Paik & Mix, 2006). More recently, this paradigm has been extended to 3-month-olds, and the results show that both relational ability and the signatures of relational learning are already present at this young age (Anderson et al., 2018). This work provides compelling evidence that the human talent for relational learning begins early.

This is not to say that infants are capable of the complexities that characterize adult relational reasoning. Adults benefit from extensive experience in the world, mature executive functioning and memory, and from another uniquely human talent – language. The studies with prelinguistic infants establish that language is not *necessary* for relational, but they do not tell us whether language can *influence* infants’ relational learning. In this paper, we explore whether language affects infant relational learning in the first year.

Gentner and colleagues have proposed that language supports relational learning and reasoning in several different ways (Gentner, 2010; Gentner & Christie, 2010). Two of these are central to the present research. First, common labels invite comparison and abstraction (Gentner, 2003; Gentner & Christie, 2008; Gentner & Namy, 1999). By giving two things the same name, we invite children to compare them; this promotes noticing common relational structure. Second, naming a concept promotes reification: learning a label for a relational structure helps to unify it and preserve it, and renders it more accessible for future use.

Christie and Gentner (2014) found evidence for both these mechanisms in research on children’s ability to notice and match the *same* relation. As baseline, they first gave children a *same-only* version of the Relational-Match-to-Sample (RMTS) task. In this task, people are given a standard (e.g., AA) and must say which alternative it matches—BB (the correct relational match) or CD. When children were given eight RMTS trials, with neither practice nor feedback, 4-year-olds passed,² but 3-year-olds failed. Christie and Gentner (2014) then asked whether learning labels for *same* and *different* would improve performance. They taught 3-year-olds to label pairs as *same* or *different*, and then gave them the same RMTS task³ as before. With this training, 3-year-olds readily passed the RMTS task—evidence that learning relational labels can strengthen a relational concept. In their final study, the researchers tested the second claim—that applying common labels supports relational cognition by encouraging people to compare things that share common labels. Children were given the same RMTS task, but this time were told “This one (the standard) is a truffet. Can you show me which one of these is also a truffet?” With these instructions, 3-year-olds succeeded at choosing the relational match. Thus, relational insight can be supported by known relational labels, but also by novel ones, that invite comparison across examples.

There is also considerable evidence that, through these routes, language can help children overcome the challenge of individual object differences, which hurts the ability to perceive relations (Christie & Gentner, 2014; Loewenstein & Gentner, 2005; Rattermann & Gentner, 1998). For example, Rattermann and Gentner (1998) presented 3-year-olds with two sets of three objects that monotonically increased in size. The children were shown a search task in which they had to match the objects based on their relative size (e.g., largest to largest). The challenge was that the largest object in one set was the same size as the medium object in the other set. Children chose correctly only 32% of the time (chance performance), despite receiving corrective feedback on each trial. However, when children were taught to label the objects as “daddy”, “mommy”, and “baby”, they made successful relational matches 79% of the time. The benefits of a known relational label have been found for other familiar words and sets of words, such as “same”, “tiny/-little/big” and “top/middle/bottom” (Casasola, 2005; Christie & Gentner, 2014; Loewenstein & Gentner, 2005; Son, Smith, Goldstone, & Leslie, 2012). Positive effects of labels have also been found when a novel word is used, such as “dax” or “blicket” (Christie & Gentner, 2010, 2014; Gentner & Namy, 1999; Namy & Gentner, 2002). These findings with novel words are clear evidence that one way in which a label improves relational learning is by inviting the learner to compare. Thus it is possible that infants might benefit from hearing relational labels even before they have acquired the words for them.

The counterargument to this possibility is that relational concepts are acquired slowly by children (Gentner, 2005; Gentner, Anggoro, & Klibanoff, 2011). Specifically, words for *same* and *different* relations – our focus in these studies – are acquired fairly late, such that only 3% of 12-month-olds are reported by MB-CDI data to understand the word “same”, and the word “different” does not

² Have generalized these findings. Using the same methods as in Christie and Gentner (2014), they found that 4-year-olds can pass the RMTS task with both *same* and *different* trials.

³ The labeling task used different materials than the subsequent RMTS task, and children were not told that the two tasks were connected, nor were they encouraged to use labels during the RMTS task.

appear on these assessments at all (Frank, Braginsky, Yurovsky, & Marchman, 2016). Nonetheless, hearing *same* and *different* as labels for relational pairs could have benefits before these terms are acquired.

Experiment 1 examines the impact of relational labels on learning *same* and *different* in the first year. We test infants at 12 months of age, around the time that they are beginning to acquire somewhat abstract words (Bergelson & Swingley, 2013). We use a habituation/dishabituation paradigm based on Ferry et al. (2015), but with a key difference. In Ferry et al., infants were shown a set of individual objects before being habituated to *same* or *different* pairs – with the result that they failed to distinguish *same* and *different* in test trials containing these objects. In our current experiment, drawing on Christie and Gentner (2014)'s methods, we use this pre-task experience to test the effect of relational labels. Ferry et al. (2015) showed infants individual objects without labels prior to habituation. In Experiment 1, we showed objects in pairs with the same relational label (*same* or *different*) applied across the three pairs. If language and relational learning interact in pre-verbal infants as they do in older children, then 12-month-olds should be able to generalize the relations to all pairs at test.

A final consideration is that infants might show differences depending on whether they hear the label for *same* or for *different*. Some researchers have theorized that infants conceptualize *same* before they conceptualize *different* (Addyman & Mareschal, 2010; Hochmann, Carey, & Mehler, 2018). Other relational studies have found comparable learning advantages for infants, whether they are habituated to *same* or to *different* (Anderson et al., 2018; Ferry et al., 2015). To address the possibility of an advantage for *same*, we test for an interaction with habituation condition, asking whether 12-month-old infants will be more likely to benefit from the label for *same* than for *different*.

2. Experiment 1

Experiment 1 examined whether prior opportunity to learn the key relational label (*same* or *different*) would help 12-month-old infants notice and abstract the relation during habituation. Prior to habituation, the infants saw three pairs of objects⁴ that all shared a common relation and were labeled with a *Relational Label*. For example, infants in the *same* condition would see three *same* pairs that were each labeled “These are *same*” as they were presented. In the study room, infants were then habituated to another four pairs with the target relation (e.g., *same* if they had heard the label *same*, or *different* if they had heard *different*), presented silently over the course of 6–9 habituation trials. Finally, infants saw three types of test trials: *New* trials with previously unseen objects, *Pair Habituation* trials with objects that appeared in the habituation phase, and *Relational Label* trials with the objects that had been given a relational label before habituation. Each of these test trial types were composed of two back-to-back trials: one featuring the familiar relation (e.g., *same* if they had heard the label *same*) and one featuring the novel relation (e.g., *different* if they had heard *same*). Infants habituated to *different* were presented with an analogous procedure, where three *different* pairs were each labeled “These are *different*” and the novel relation at test was *same*. (While *same* and *different* are typically embedded in different syntactic structures – “These are the *same*” vs. “These are *different*” – we wanted to equate syntax across conditions, opting for “These are *same*.”).

If the pretrial training with labeled exemplars of the relation supports relational learning during habituation, then infants should look longer at the novel relation than the familiar relation across all test types. This includes the critical test of generalizing the relation to never-before-seen objects in the *New* trials. If labeling relations helps infants detect those relations in the subsequent habituation experience, then infants should also look longer at the novel relation compared to the familiar one in the *Relational Label* and *Pair Habituation* trials, where novel and familiar relations are made of objects that appeared prior to test. In the *Pair Habituation* trials, both novel and familiar relations are made of objects that appeared in the habituation phase; but the familiar pair is an exact pair from habituation, while the novel pair that mixes objects from two other habituation trials. In the *Relational Label* trials, both novel and familiar relations are made of objects that appeared during labeling, but the familiar pair is one of the exact pairs that was labeled, while the novel pair mixes objects from the other two labeled pairs.

As mentioned earlier, there may be an interaction with habituation condition, such that *same* is easier to learn than *different* (Hochmann et al., 2018). If this is the case, infants in the *same* condition should discriminate novel and familiar trials, while infants in the *different* condition may not.

3. Methods

3.1. Participants

The participants were 64 12-month-old infants, ranging from 10 months, 29 days to 13 months 3 days (33 females; $M_{Age} = 12$ months, 0 days). This sample size was chosen because it was the same as that of Ferry et al. (2015), and a power analysis confirmed that even with the smallest meaningful effect size from that study ($\eta_p^2 = .12$; $f = 0.369$), this sample size would result in at least 95% power for a repeated measures ANOVA testing the within-subjects factor of relation (2 measures) with 8 between-subject cells (Condition $\times$ Sex $\times$ Order) (G*Power v3.1.9). Half of these infants were assigned to the *same* condition and half to the *different* condition. Seven additional infants were excluded: six for fussiness and one for a bowel movement during test (see Section 3.5 for how fussiness was defined).

Parents of infants were recruited through online ads and word of mouth. Parents who agreed to their infants' participation were

⁴ As in Christie and Gentner's (2014) Experiment 3, multiple examples of the relation were used in the pretraining with the relational label. However, due to the constraints of infants' attention span, the infants received only three examples. We return to this issue in the discussion.

provided informed consent before the experiment and given \$20 as compensation. The race, ethnicity and parental education level of the sample is described in the table below (Table 1).

3.2. Apparatus

Parents sat in a chair with infants on their lap facing a wooden puppet stage that displayed all stimuli. The parents were asked to refrain from interacting with the infant during the experiment and to close their eyes during the test trials. The stage measured 243.5 cm high, 128 cm wide, and 61 cm deep. The opening in the front of the stage that displayed the objects was 93 cm above the floor, 61 cm high, and 106 cm wide. The back wall had two rectangular openings with cloth fringe over the openings that allowed the experimenter to manipulate the objects. A screen that covered the infants' view of the stage was raised and lowered between trials. The MATLAB program Baby Looking Time (BLT) was used to record looking times for habituation and test trials during the experiment (Chang, Wang, & Way, 2018).

3.3. Stimuli

The stimuli consisted of 18 three-dimensional objects (see Fig. 1). During the experiment, each pair of objects was placed on the puppet stage on a 26.5×15.5 cm cardboard tray that was covered with contact paper. Six of the objects were presented in the waiting room beforehand in the context of *same* or *different* pairs. Four additional *same* or *different* pairs were shown in the habituation phase. Finally, 12 objects appeared in pairs during the test phase. Of these, four objects had been labeled in pairs in the waiting room, four objects had been shown in pairs in habituation, and four objects had not been seen before test.

3.4. Procedure

The experiment consisted of three sequential parts:

3.4.1. Relations are labeled

The experimenter showed three pairs of objects to the infant in the waiting room prior to the experiments, while labeling the relations they instantiated. These were either three pairs of two identical objects (for the *same* condition) or three pairs of two non-identical objects (for the *different* condition). During each pair presentation, the experimenter held the object in front of the infant for 5 s and said to the infant, "These are *same*!" or "These are *different*!" Identical sets of objects were used for both conditions: two blue aardvarks, two white espresso cups and two pink and green foam towers (see Fig. 1-A for configurations).

3.4.2. Pair habituation trials

When the screen was raised at the start of every trial, a pair of objects rested on the cardboard tray on the stage. To engage infants' attention, the experimenter held and moved the objects throughout the trial. The experimenter grasped one object in each hand and raised the objects straight up (1 s), tilted them to the left (1 s), returned them to the center (1 s), tilted them to the right (1 s), returned them to the center (1 s), returned them to the tray (1 s), and paused on the tray (2 s). This 8-s cycle repeated continuously until the trial ended. In the *same* condition, infants saw habituation trials in which the pairs of objects were the *same* (see Fig. 1-B). In the *different* condition, infants saw habituation trials in which the pairs of objects were *different*. The number of habituation trials was infant-controlled (see Section 3.5 for the criterion) and ranged from 6 to 9 trials.

3.4.3. Test trials

For each test trial type, infants were presented with a novel relation, followed by a familiar relation (or vice versa), summing to six test trials (see Fig. 1-C). In each test trial, infants viewed one pair of objects, presented in the same motion pattern as in the habituation trials, while their looking time was recorded. The three kinds of test trials were (a) objects that had not been seen before test (*New*); (b) objects that the infant had heard labeled in the waiting room (*Relational Label*); and (c) objects that the infant had seen presented in the pair habituation trials (*Pair Habituation*). There were three trial orders (abc, cab, bca) which were counterbalanced across infants.

Table 1
Participant demographic information.

Ethnicity	%	Race	%	Education	Maternal%	Paternal%
Hispanic	10	African American	8	Some high school	0	0
Non-Hispanic	83	Asian/Hawaiian/Pacific Islander	1	High school diploma	0	0
		American Indian	0	Some college	5	10
		White	75	College degree	91	84
		Multiracial	10			
No response	7	No response	6	No response	4	6

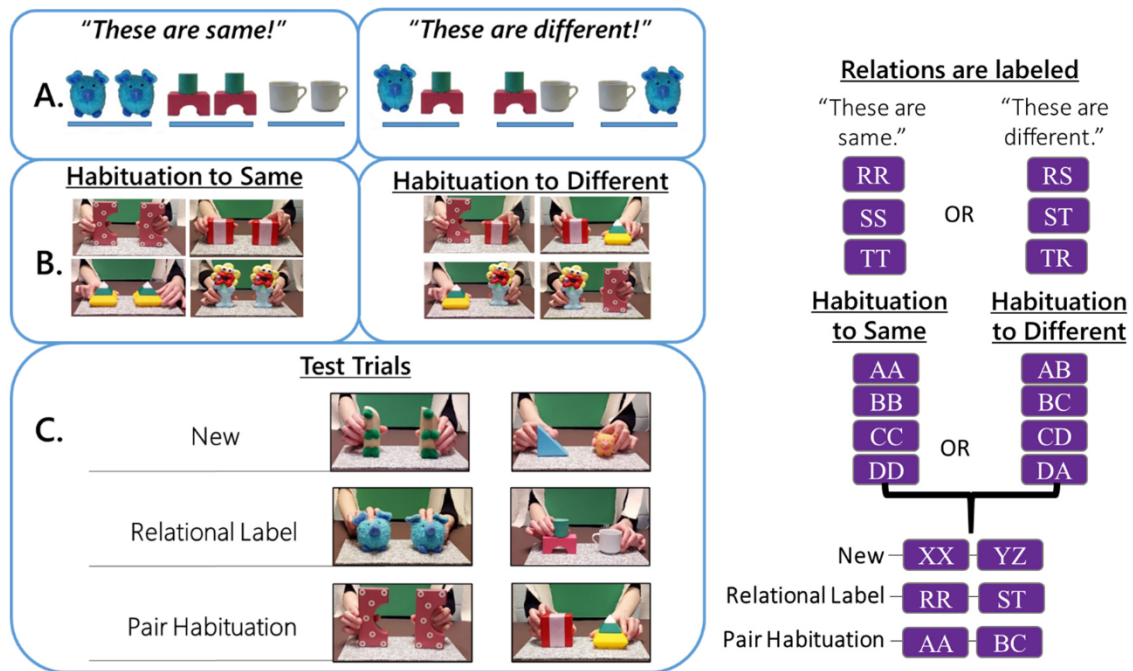


Fig. 1. Pictures of the stimuli used in Experiment 1 and a schematic of the procedure. (A) Infants were presented with three pairs before the experiment, labeled appropriately as same or different. (B) Infants were habituated to four same or different pairs. (C) Infants saw three types of test trials: *New*, *Relational Label*, and *Pair Habituation*. The order of trials was counterbalanced across participants.

3.5. Coding

There was a small hole in the front face of the stage containing a camera that captured a video image of the infant's face. While the experimenter conducted habituation and test trials in the room with the infants, two research assistants in a separate room viewed a video of the infants' face and coded the infants' visual fixations online as either on- or off-target. Each researcher depressed a computer button when the infant attended to the events on stage and released the button when the infant looked away. Each trial ended when the infant either looked away for 2 consecutive seconds after having looked at the event for at least 2 s or looked at the event for 60 cumulative seconds without looking away for 2 consecutive seconds. This 2-s criterion has been standard practice in the Hespos lab for more than 10 years (Hespos, Ferry, & Rips, 2009; Hespos, Grossman, & Saylor, 2010). In this study, we also wished to maintain the same procedures as had been used in Ferry et al. (2015), to preserve comparability between their results and ours.

The BLT program indicated the end of the trial with a beep, signaling to the experimenter to lower the screen and move to the next trial. After each test trial, research assistants also checked one or more boxes to indicate the behavioral state of the infant on the preceding trial: sleepy, quiet and alert, active, fussy or crying. Coders also noted any breaks and their length. As noted above, if two coders independently judged the infant's state as fussy, crying, or falling asleep for more than half the test trials, the infant's data was excluded from the analysis. The coders were blind to the condition and the trial order. Interobserver agreement was measured for all infants and averaged 91%. The Intraclass Correlation Coefficient for a fixed set of raters (ICC3) was .91 with a 95% confidence interval from .88 to .93, $F(40, 233) = 11$, $p < .001$. Our looking time data were skewed towards the lower bound and significantly deviated from a normal distribution per the Shapiro-Wilk test, $W = 0.82$, $p < .001$. Therefore, we performed parametric tests on log-transformed data, following recommendations outlined by Csibra, Hernik, Mascaro, Tatone, and Lengyel (2016). The log-transformed data did not differ from the normal distribution, $W = 0.98$, $p = .24$.

4. Results

The key question was whether relational labels would facilitate learning in 12-month-olds. The answer is no: there was no clear evidence of learning for any of the test trial types. Overall, infants looked an average of 19.39 s ($SD = 10.37$) at pairs of objects representing novel relations and 16.72 s ($SD = 8.6$) at pairs representing familiar relations. An omnibus repeated-measures ANOVA tested the within-subjects variables of Relation (novel vs. familiar) and Test Type (*Pair Habituation*, *Relational Label*, and *New*) and the between-subjects variables of Condition (habituation to *same* vs. *different*), Test Order (novel relation or familiar relation first), and Sex (male vs. female), and failed to show a main effect of relation, $F(1,53) = 3.38$, $p = .072$. There was no interaction between relation and test type, $F(1,53) < 1$, $p = .823$, nor an interaction between relation and condition, $F(1,53) = 1.61$, $p = .211$. However, there was a four-way interaction between Relation, Test Type, Condition and Order, $F(1,53) = 3.38$, $p = .038$. To better understand this interaction, we conducted separate ANOVAs for each of the test types, including the factors of condition and order.

We first conducted a repeated-measures ANOVA on the *New* test trial type, looking at the within-subjects variable of Relation (novel vs. familiar) and the between-subjects variables of Condition (habituation to *same* vs. *different*), and Test Order (novel relation or familiar relation first). Averaging across groups, infants looked 21.10 s ($SD = 16.16$) at novel relations and 16.95 s ($SD = 11.85$) at familiar ones. There was no main effect of relation, $F(1, 60) = 2.80$, $p = .10$. A binomial comparison also reflects the lack of discrimination: only 37 infants out of 64 looked longer at the novel relation in the *New* trials, $p = .130$. However, there was an interaction between relation and condition, $F(1, 60) = 8.44$, $p = .005$, $\eta_p^2 = .12$: infants in the *same* condition looked longer at the novel relation than the familiar one ($M_{Novel} = 24.13$ s; $M_{Familiar} = 15.49$ s, $t(32) = 2.67$, $p = .012$), while infants in the *different* condition looked equally between them ($M_{Novel} = 17.89$ s; $M_{Familiar} = 18.51$ s, $t(30) = -0.93$, $p = .362$), see Fig. 2. A binomial comparison also suggested that infants' relational generalization fared better in the *same* condition: 22 of the 33 in the *same* condition looked longer at the novel relation than the familiar one, $p = .04$, while only 16 of 31 infants in the *different* condition did so, $p = .64$. There was also an interaction between relation and test order, such that infants who saw novel relations first looked during test longer at the novel relation than the familiar one ($M_{Novel} = 21.24$ s; $M_{Familiar} = 14.49$ s), while infants who saw the familiar relation first during test looked roughly equally between them one ($M_{Novel} = 20.97$ s; $M_{Familiar} = 19.42$ s), $F(1, 60) = 5.49$, $p = .02$, $\eta_p^2 = .08$. Finally, there was a three-way interaction between relation, condition and test order. The infants who saw the novel relation first in the *same* condition showed greater looking differences between novel and familiar relation than did infants in any of the other cells, $F(1, 60) = 5.70$, $p = .02$, $\eta_p^2 = .09$. There was no main effect of condition on overall looking, $F(1, 60) < 1$, $p = .64$, nor a main effect of order, $F(1, 60) < 1$, $p = .70$. There were no other significant main effects or interactions.

Next, we conducted a repeated-measures ANOVA on the *Relational Label* test trial type, looking at the within-subjects variable of

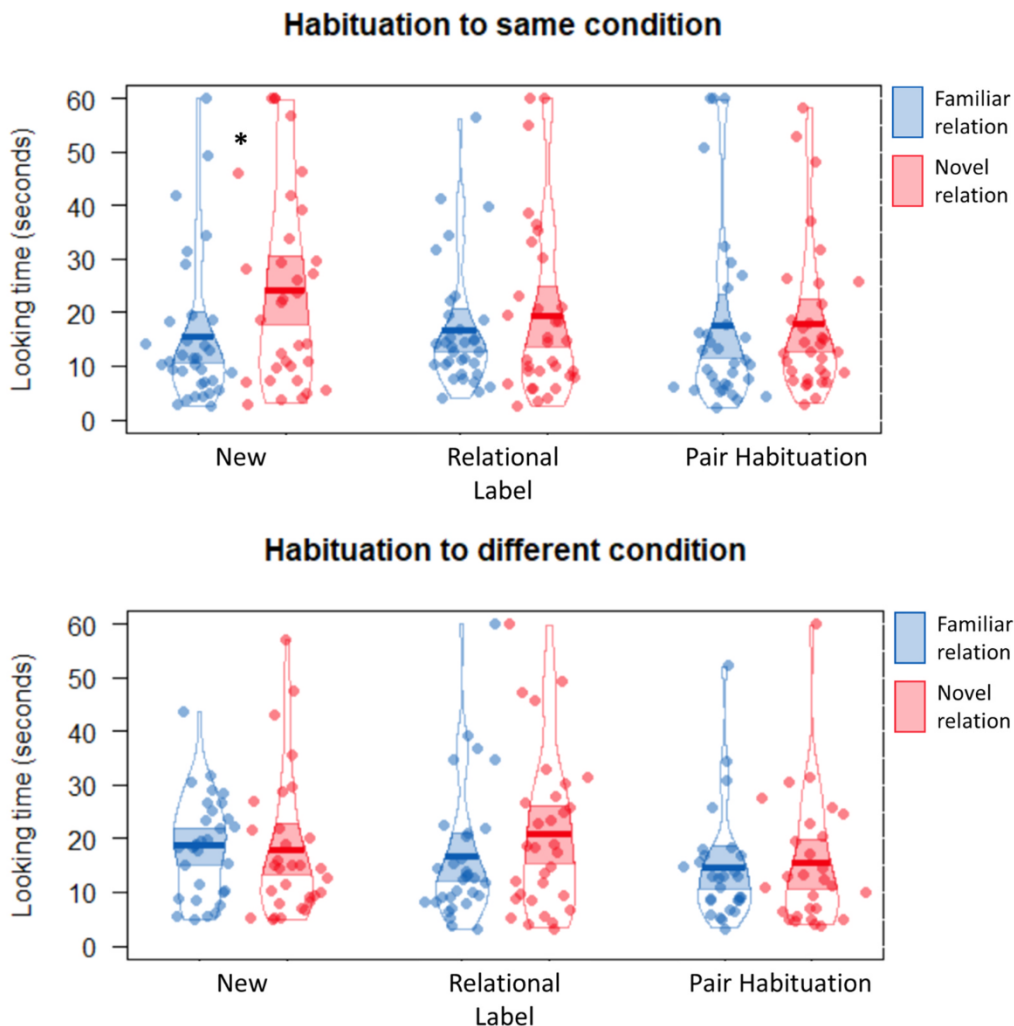


Fig. 2. Experiment 1 test trial looking times for 12-month-olds who heard relational labels, separated by habituation condition (*same* vs. *different*) and by test trial type (*New*, *Relational Label*, and *Pair Habituation*). The thick central line in each box is the mean, and the upper and lower shaded portions represent the 95% Confidence Intervals (CIs) for this mean. Dots indicate the raw data points. The width of the bean indicates the density of the data distribution at a looking time value. * indicates $p < .05$.

Relation (novel vs. familiar) and the between-subjects variables of Condition (habituation to *same* vs. *different*), and Test Order (novel relation or familiar relation first). Averaging across groups, infants looked 20.32 s ($SD = 15.31$) at novel relations and 17.02 s ($SD = 12.33$) at familiar ones, and there was no main effect of relation, $F(1, 60) < 1, p = .354$. Again, the binomial comparison reflects this pattern: only 34 infants out of 64 looked longer at the novel relation in the *Relational Label* trials, $p = .354$. There was no interaction between relation and condition, $F(1, 60) < 1, p = .409$. A binomial comparison also suggested that infants performed similarly in both conditions: 18 of the 33 in the *same* condition looked longer at the novel relation than the familiar one, $p = .364$, and only 16 of 31 infants in the *different* condition did so, $p = .50$. There were no other significant main effects or interactions.

Finally, we conducted a third repeated-measures ANOVA on the *Pair Habituation* test trial type, looking at the within-subjects variable of Relation (novel vs. familiar) and the between-subjects variables of Condition (habituation to *same* vs. *different*), and Test Order (novel relation or familiar relation first). Infants looked near equally at novel ($M = 16.60$ s, $SD = 13.14$) and familiar relations ($M = 16.05$ s, $SD = 14.34$), and there was no main effect of relation, $F(1, 57) < 1, p = .636$. Again, the binomial comparison reflects this: only 33 of 61 infants looked longer at the novel relation in the *Pair Habituation* trials, $p = .304$. There was also no interaction with condition, $F(1, 57) < 1, p = .508$. A binomial comparison also suggested that infants performed similarly in both conditions: 21 of the 33 in the *same* condition looked longer at the novel relation than the familiar one, $p = .081$, and only 12 of 28 infants in the *different* condition did so, $p = .828$. There were no other significant main effects or interactions.⁵

4.1. Comparison to Ferry et al. (2015)

To understand the effect of labeling the relation compared to no labeling, we compared the 12-month-old infants in the Experiment 1 to the 7- and 9-month-olds in Ferry et al. (2015) for each test type.⁶

Did labeling the relation improve generalization compared to Ferry et al. (2015)? The answer was no. We conducted a repeated measures ANOVA on the *New* trials, examining the within-subject factor of Relation and the between-subject factors of Experiment (Ferry et al., 2015 vs. current study) and Condition (habituation to *same* vs. *different*). This showed a main effect of Relation, overall, $F(1, 124) = 10.84, p = .001, \eta_p^2 = .08$. This effect of Relation did not depend on Experiment: although infants looked significantly longer at the novel relation in Ferry et al. (2015) ($M_{\text{Novel}} = 18.91$ s; $M_{\text{Fam}} = 13.09$ s) and not in this experiment ($M_{\text{Novel}} = 21.10$ s; $M_{\text{Fam}} = 16.95$), there was not a significant interaction, $F(1, 124) = 1.16, p = .284$. The effect of Relation did not interact with Condition, either, $F(1, 124) = 2.53, p = .114$. However, there was a three-way Relation $\times$ Experiment $\times$ Condition interaction, in that infants in Ferry et al. (2015) generalized in both the *same* and *different* conditions, while infants in the current experiment only generalized when habituated to *same* $F(1, 124) = 5.45, p = .021, \eta_p^2 = .04$. There were no other main effects.

The relational label also did not have a positive impact on the *Relational Label* trials, compared to the unlabeled *Salient Object* trials in Ferry et al. (2015). A repeated measures ANOVA on the *Relational Label/Salient Object* trials, examining the within-subject factor of Relation and the between-subject factors of Experiment (Ferry et al., 2015 vs. current study) and Condition (habituation to *same* vs. *different*) showed no main effect of Relation, $F(1, 124) < 1, p = .820$. The lack of generalization did not interact with Experiment. That is, infants in Ferry et al. (2015) looked near equally between relations ($M_{\text{Novel}} = 17.74$ s; $M_{\text{Fam}} = 17.23$ s) as well as with labeled pairs in the current experiment ($M_{\text{Novel}} = 19.90$ s; $M_{\text{Fam}} = 16.61$), $F(1, 124) = 1.69, p = .197$. There was no interaction between Relation and Condition, $F(1, 124) = 1.13, p = .290$, nor was there a three-way interaction between these factors and Experiment, $F(1, 124) < 1, p = .739$. There were no main effects.

Finally we performed a repeated measures ANOVA on the *Pair Habituation* trials, examining the within-subject factor of Relation and the between-subject factors of Experiment and Condition. The ANOVA showed a main effect of Relation, $F(1, 121) = 9.36, p = .003, \eta_p^2 = .07$, and an interaction with Experiment, $F(1, 121) = 5.05, p = .026, \eta_p^2 = .04$. Infants in Ferry et al. (2015), who did not hear labels, discriminated between pairs from habituation ($M_{\text{Novel}} = 18.31$ s; $M_{\text{Fam}} = 13.21$ s), while infants in Experiment 1, who heard labels, looked equally between relations ($M_{\text{Novel}} = 16.60$ s; $M_{\text{Fam}} = 16.05$ s). There was no interaction between Relation and Condition, $F(1, 121) < 1, p = .95$, nor was there a three-way interaction between these factors and Experiment, $F(1, 121) < 1, p = .36$.

4.2. Habituation patterns

Finally, we examined looking time during the habituation phase. Some theories that argue that *same* should be easier to learn than *different* (Addyman & Mareschal, 2010; Hochmann et al., 2018). In the context of this experiment, this would suggest that the habituation decline between conditions should be commensurate to the ease of learning *same* compared to *different*. A repeated measures ANOVA examined looking time by the within-subjects factor of Habituation Trial (first three through last three) and the between-subjects factor of Condition (habituation to *same* vs. *different*), and revealed a significant decline in looking across conditions, $F(5, 310) = 57.30, p < .001, \eta_p^2 = .43$ (see Fig. 3). On average, infants looked 46.96 s on their first habituation trial and 15.66 s on their last trial. However, there was not a significant interaction between Habituation Trial and Condition, $F(5, 310) = 1.042, p < .382$, nor a main effect of Condition, $F(1, 62) < 1, p < .536$. Additionally, Condition (*same* or *different*) did not predict total looking time

⁵ A pilot experiment implemented this procedure with 7- and 9-month-olds ($n = 53$) and found null results for all test trials: *New*, $t(52) = 3.472, p = .068$; *Relational Label*, $t(52) < 1, p = .987$; and *Habituation*, $t(52) < 1, p = .368$.

⁶ There were minor differences between pair habituation trials in the two experiments. Infants in Ferry et al. (2015) saw objects presented in both in the waiting room and then in habituation pairs before they finally reappeared in test trials, while infants in Experiment 1 only saw the objects in habituation before they reappeared at test.

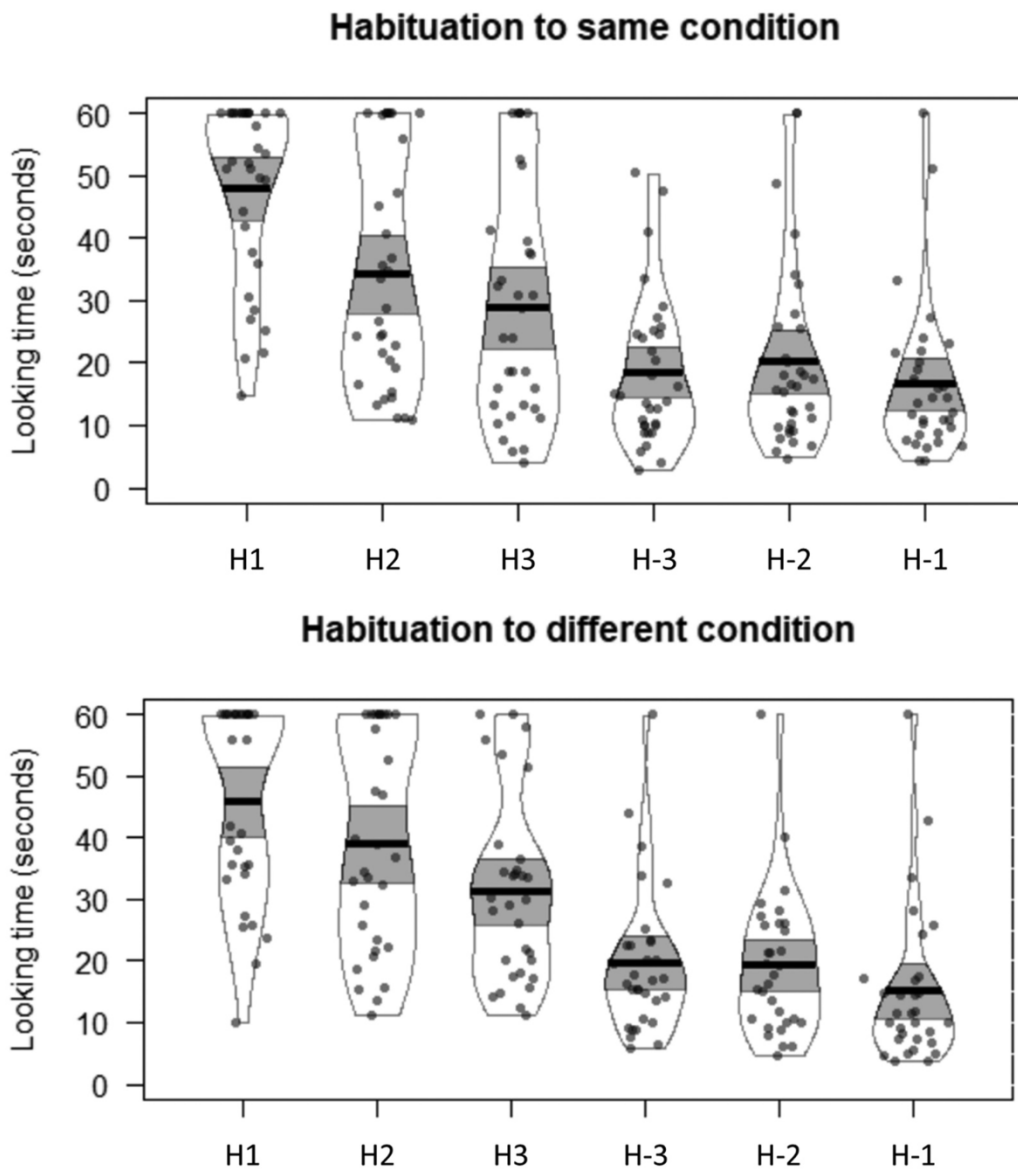


Fig. 3. Habituation trial looking times. Because the number of habituation trials was infant-controlled, ranging from 6 to 9 trials, the graphs represent the first three (H1–H3) and the last three trials (H-3, H-2, H-1). The thick central line in each box is the mean, and the upper and lower shaded portions represent the 95% Confidence Intervals (CIs) for this mean. Dots indicate the raw data points. The width of the bean indicates the density of data points at a looking time value.

during habituation, $F(1, 62) < 1$, $p = .743$, nor how many trials it took to reach the habituation criterion, $X^2(3, N = 64) = 2.32$, $p = .509$, nor whether infants reached the habituation criterion, $X^2(1, N = 64) < 1$, $p = .47$. Additionally, infants' looking times on the first habituation trial was 47.87 s in the *same* condition and 45.74 s in the *different* condition, suggesting that baseline looking time for *same* relations was not greater than for *different*, $t(62) = 0.57$, $p = .57$.

5. Discussion

In Experiment 1, hearing the relevant relational label applied to three pairs just prior to the habituation did not improve 12-month-olds' learning *same* or *different*. Specifically, labels did not help infants discriminate between relations on the *Relational Label* pairs. Moreover, infants across conditions failed to distinguish *same* and *different* relations during *Pair Habituation* test trials, constituting a significant difference from performance in the [Ferry et al. \(2015\)](#) no-label paradigm. This failure is striking because differentiating a pair seen in habituation (e.g., AA) from a novel pair made of habituation objects (e.g., BC from BB and CC) should not require any relational learning—it simply requires memory retrieval.

Infants in Experiment 1 did not differ from Ferry et al.'s (2015) no-label paradigm in their ability to generalize to *New* pairs. However, unlike in Ferry et al. (2015), infants in Experiment 1 who habituated to *same* fared better than those habituated with *different*, generalizing this relation to *New* test trials. We had considered the possibility that a *same* label might be easier to learn than a *different* label, as has been found in other infant studies (Addyman & Mareschal, 2010; Hochmann et al., 2018) and based on MB-CDI data showing that *same* is understood and produced earlier (Frank et al., 2016). However, if the relative ease of the *same* label is what allowed infants to generalize in this condition, then infants should also have discriminated between relations on all trials. Instead, infants in the *same* condition looked equally at novel and familiar relations in the *Relational Label* and *Pair Habituation* test trials. This lack of discrimination suggests that the relational label did not provide any benefits, even to infants in the *same* condition.

Previously, we asked if relational labels might improve relational learning even in infancy. Infants' performance in Experiment 1 argues that this was not the case. How do we reconcile this with the literature showing benefits of relational language for preschool age children? The most obvious answer is age. While infants may understand some object words like "bottle" and "apple" at 9 months, experiments suggest that they do not seem to understand more relational terms like "all gone" until 12 months (Bergelson & Swingley, 2012, 2013). Parent reports of infants' comprehension show similar patterns, with names for objects or specific individuals (e.g., "mama" and "dada") being reported before verbs and other relational terms (Frank et al., 2016). (Gentner, 1982; Gentner & Boroditsky, 2009) have suggested that the acquisition pattern reflects the greater difficulty of discerning relational concepts in the environment. Thus, even if a relational label increased the opportunity for comparison, infants with little relational language experience may struggle to map labels onto relations without greater support. In our experiment, infants may have been unable to discern the concept from the three examples they received.

The syntax of how we labeled *same* could also have supported an object category mapping over a relational one. While we used "these are *same*" to keep the syntax identical between the *same* and *different* conditions, this phrasing is not grammatical in English. Instead, this syntax matches that of plural noun names, e.g., "These are cups". The work on syntactic bootstrapping has shown that the surrounding syntax is an important clue to the relational structure of novel verbs, i.e., whether they take one or two arguments (Fisher, 1996; Gleitman, 1990; Gleitman, Cassidy, Nappa, Papafragou & Trueswell, 2005). A similar phenomenon could be at play with "these are the *same*" with the article after the plural verb "are" signaling that the label "same" is not an object name ("cups") or an object property ("blue").

In Experiment 1, we hypothesized that labeling relations would aid relational abstraction. We failed to find evidence for this supposition. In Experiment 2, we ask whether labeling distinct objects within and between relational pairs would hinder relational learning. There is ample evidence that object words are a part of infants' early vocabulary (e.g., Bergelson & Swingley, 2012) and that drawing attention to objects interferes with infants' relational learning (Anderson et al., 2018; Ferry et al., 2015). Just as the preschool literature indicates that common labels, and prior learned relations, should help relational learning, it also offers evidence that object labels can interfere with relational processing. While four-year-old children can typically pass a Relational Match to Sample (RMTS) task, they fail this task if it is preceded by an object naming task—even though the named objects never appeared in the RMTS task (Hoyos, Shao, & Gentner, 2016). Hoyos et al. (2016) speculated that the object-naming task focuses children on the individual characteristics of objects in a scene, to the detriment of their relational processing. In Experiment 2, we ask if labeling objects disrupt relational learning in a predictable way by 12 months.

6. Experiment 2

Experiment 2 examined the influence of object labels on 12-month-old infants' relational learning. Prior to habituation, the infants saw three pairs of objects. In contrast to Experiment 1, these objects were shown individually instead of in pairs, and each object received its own individual label (see Fig. 4). The labels were English words whose content matched the items presented (e.g., *cup*, *tower*). The sequence of objects and their labels was presented twice. Next, in the study room, infants were habituated to four pairs that shared a common relation (e.g., all *same* or all *different*). These were presented silently over the course of 6–9 habituation trials. Finally, infants saw four pairs of test trials: *New* trials with previously unseen objects, *Object Label* trials with the objects that had been given a label in the beginning, *Pair Habituation* trials with objects that appeared in the habituation phase, and *Memory Check* trials. In the *Memory Check* trials, the familiar test pair had been seen in habituation, but neither the novel pair's relation nor the objects that instantiated it had been seen before. In contrast, in *Pair Habituation* trials, all of the objects had been seen in habituation, but the novel test pair presented these objects in the context of a novel relation. The *Memory Check* trial was added as an easier version of the *Pair Habituation* trial in response to the widespread failures to discriminate relations in Experiment 1. Each of the four test trial types included two sequential trials: one featuring the familiar relation (e.g., *same* if they had learned *same*) and one featuring the novel relation (e.g., *same* if they had learned *different*).

If labeling individual objects disrupts relational learning, infants should look equally at novel and familiar relations in both the *New* trials and the *Object Label* trials. In the *Pair Habituation* and the *Memory Check* trials, infants only need to remember the exact pairs that they saw in the habituation phase to discriminate the relation. They should look longer at the novel relation in both these trials, regardless of whether they generalize the relation in other trials.

7. Methods

7.1. Participants

Participants were 48 infants, ranging from 11 months, 13 days to 12 months 15 days (18 female, $M_{Age} = 12$ months, 3 days). This

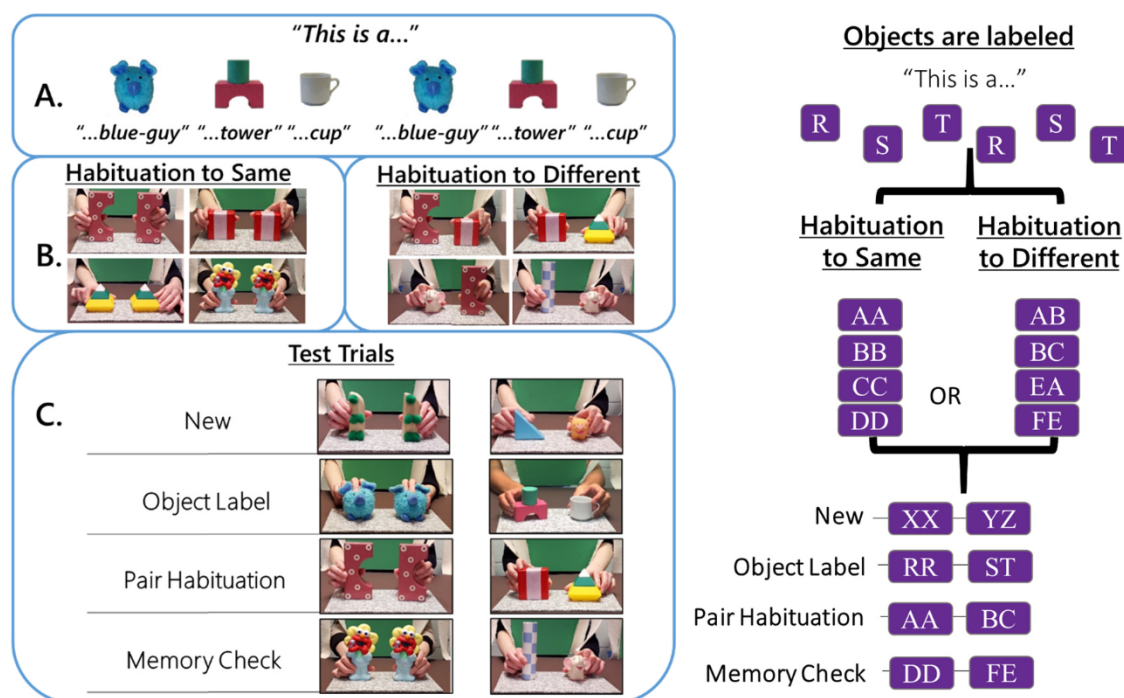


Fig. 4. (A) Infants heard individual objects labeled before the experiment. (B) Infants habituated to four *same* or *different* pairs. (C) Infants saw three types of test trials: *New*, *Object Label*, *Pair Habituation*, and *Memory Check*. The order of trials was counterbalanced across participants.

Table 2
Participant demographic information.

Ethnicity	%	Race	%	Education	Maternal%	Paternal%
Hispanic	10	African American	12	Some high school	0	0
Non-Hispanic	83	Asian/Hawaiian/Pacific Islander	6	High school diploma	0	2
		American Indian	0	Some college	10	12
		White	65	College degree	90	87
		Multiracial	17			
No response	0	No response	0	No response	0	0

sample was based on a power analysis of the Relation $\times$ Condition effect for *New* trials in Experiment 1 ($\eta_p^2 = .12$; $f = 0.369$), for a repeated measures ANOVA testing the within-subjects factor of relation (2 measures) with 8 between-subject cells (Condition $\times$ Sex $\times$ Order) (G*Power v3.1.9). Half of the infants were assigned to the *same* condition, and half to the *different* condition. An additional three infants were run but were excluded because of fussiness for over half of the test trials. Exclusion criteria was the same as in Experiment 1, as were recruitment and compensation. The demographic details of the sample are listed below (see Table 2).

7.2. Apparatus

The stage was the same as in Experiment 1. The stimuli were 20 three-dimensional objects. The objects were similar to those in Experiment 1, except that to create the *Memory Check* trials, two new objects were introduced: a white elephant and a blue-checked cylinder.

7.3. Design and procedure

The design and procedure were modeled on Experiment 1 with the following differences. The object experience phase differed in that the objects were labeled individually. Specifically, the experimenter held up each individual object and said to the infant, "This is a (cup/tower/blue-guy)!" using the appropriate label. The test trials were the same as in Experiment 1, except the *Object Label* pairs at test were composed of objects that have been labeled individually. Additionally, a fourth pair of test trials was added: *Memory Check* trials, where a habituation pair (familiar objects and familiar relation) was compared to a novel relation made of new objects.

7.4. Coding

Coding was the same as in Experiment 1. The coders were blind to the condition and the trial order. Interobserver agreement was measured for all infants and averaged 93%. The Intraclass Correlation Coefficient for a fixed set of raters (ICC3k) was .99 with a 95% confidence interval from .98 to .99, $F(47, 383) = 103$, $p < .001$. Our looking time data skewed towards the lower bound and significantly deviated from a normal distribution per the Shapiro-Wilk test, $W = 0.84$, $p < .001$, so we performed parametric tests on log-transformed data, as in Experiment 1, and the transformed data did not differ, $W = 0.98$, $p = .24$.

8. Results

The key question was whether hearing prior labels for individual objects would hurt relational learning in 12-month-olds. The results show, first, that infants could still discriminate relations overall. An omnibus repeated-measures ANOVA on the within-subjects variables of Relation (novel vs. familiar) and Test Type (pair habituation, object labels, and new) and the between-subjects variables of Condition (habituation to *same* vs. *different*), Test Order (novel relation or familiar relation first), and Sex (male vs. female) showed a main effect of relation, $F(1, 34) = 8.42$, $p = .006$, $\eta_p^2 = .198$. Averaging across test trials and other factors, infants looked 18.52 s ($SD = 14.25$) at novel relations and 16.33 s ($SD = 12.96$) at familiar ones. There was also a main effect of test type, $F(2, 68) = 4.26$, $p = .018$, $\eta_p^2 = .111$, in that infant had shorter looking times during *Pair Habituation* trials ($M = 14.45$, $SD = 11.18$) than during *New* trials ($M = 16.91$, $SD = 14.12$) or *Object Label* trials ($M = 17.75$, $SD = 13.15$), which is to be expected with a habituation paradigm.

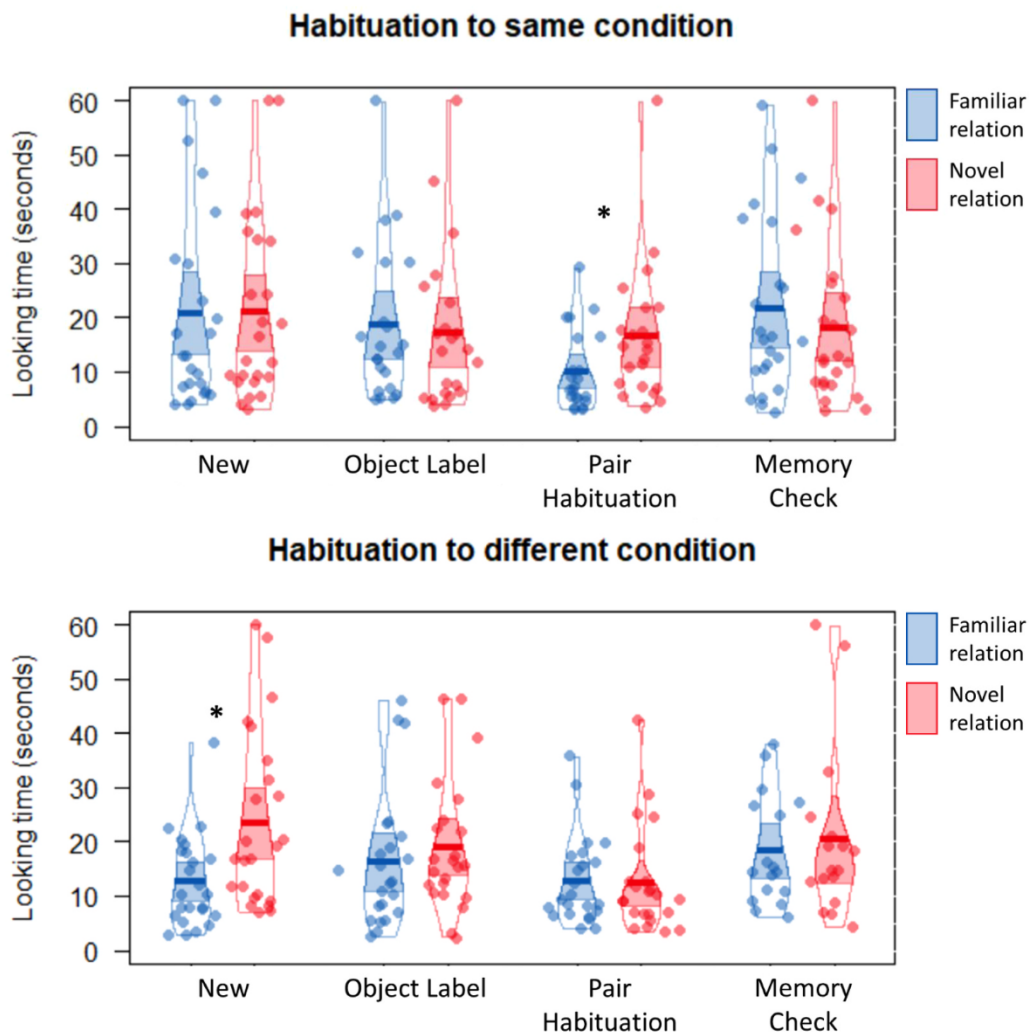


Fig. 5. Experiment 2 test trial looking times for 12-month-olds who heard individual object labels, separated by test type (*New*, *Object Label*, *Pair Habituation* and *Memory Check*). The thick central line in each box is the mean, and the upper and lower shaded portions represent the 95% Confidence Intervals (CIs) for this mean. Dots indicate the raw data points. The width of the bean indicates the density of the data distribution at a looking time value. * indicates $p < .05$.

While there was no interaction of relation and test type, $F(2, 68) = 1.89, p = .159$, there was a relation by test type by condition interaction, $F(2, 68) = 5.61, p = .006, \eta_p^2 = .142$, where infants looked longer at the *same* relations during the *Object Label* test, whether they had habituated to *same* or *different*. To better understand this interaction, we conducted separate ANOVAs for each of the test types, including the factor of condition. There were no between-subject effects and no other interactions.

The critical test for relational learning is generalization to never-before-seen objects, so we conducted a repeated-measures ANOVA on the *New* test trial type, looking at the within-subjects variable of Relation (novel vs. familiar) and the between-subjects variables of Condition (habituation to *same* vs. *different*). Averaging across groups, infants looked 22.21 s ($SD = 16.31$) at novel relations and 16.69 s ($SD = 14.64$) at familiar ones, resulting in a main effect of relation, $F(1, 46) = 8.08, p = .007, \eta_p^2 = .149$ (see Fig. 5). A binomial comparison reflected the effect of relation, with 31 of the 48 infants looking longer at the novel relation on this trial type, $p = .059$. There was a relation by condition interaction: $F(1, 46) = 5.57, p = .023, \eta_p^2 = .108$. Infants generalized the relation in the *different* condition ($M_{\text{Novel}} = 23.48$ s; $M_{\text{Familiar}} = 12.66$ s, $t(23) = 3.60, p = .002$), but not in the *same* condition ($M_{\text{Novel}} = 20.94$ s; $M_{\text{Familiar}} = 20.73$ s, $t(23) = 0.35, p = .73$). A binomial comparison also reflects this: 19 of the 24 in the *different* condition looked longer at the novel relation than the familiar one, $p = .007$, while 12 of 24 infants in the *same* condition did so, $p = 1$. There was no main effect of condition, $F(1, 46) < 1, p = .72$.

Next, we conducted a repeated-measures ANOVA on the *Object Label* test trial type, looking at the within-subjects variable of Relation (novel vs. familiar) and the between-subjects variables of Condition (habituation to *same* vs. *different*). As would be expected if object labels hampered relational focus, there was no difference in looking time between novel relations ($M = 18.11$ s; $SD = 13.24$) and familiar ones ($M = 17.39$ s; $SD = 13.19$), and there was no main effect of relation, $F(1, 43) < 1, p = .75$. A binomial comparison reflected this lack of discrimination, with 21 of the 45 infants with usable data for this trial type looking longer at the novel relation on this trial type, $p = .766$. There was no interaction with condition indicating an advantage for *same* over *different*, $F(1, 43) = 1.84, p = .18$. There was no main effect of condition, $F(1, 43) < 1, p = .89$.

A repeated-measures ANOVA on the *Pair Habituation* test trial type, looking at the within-subjects variable of Relation (novel vs. familiar) and the between-subjects variables of Condition (habituation to *same* vs. *different*), showed a nonsignificant trend towards an effect of relation in the predicted direction, $F(1, 43) = 3.89, p = .055$. Infants looked an average of 14.45 s ($SD = 11.18$) at novel relations and 11.48 s ($SD = 7.81$) at familiar relations. A binomial comparison reflected an effect of relation, with 30 of the 45 infants with usable data for this trial type looking longer at the novel relation on this trial type, $p = .036$. Again there was a significant interaction with condition, this time reflecting an advantage for *same*, $F(1, 43) = 7.59, p = .009, \eta_p^2 = .15$: infants showed discrimination in the *same* condition ($M_{\text{Novel}} = 16.54$ s; $M_{\text{Familiar}} = 10.12$ s, $t(21) = 3.08, p = .006$) but not in the *different* condition ($M_{\text{Novel}} = 12.44$ s; $M_{\text{Familiar}} = 12.77$ s, $t(22) = -0.61, p = .551$). Similarly, a binomial comparison showed that 18 of 22 infants discriminated the novel and familiar pairs in the *same* condition, $p = .004$, but only 12 of 23 discriminated in the *different* condition, $p = 1$. There was no main effect of condition, $F(1, 43) < 1, p = .99$.⁷

Experiment 2 also included a *Memory Check* test trial, a simplified version of the *Pair Habituation* trial, where the novel pair was also made of novel objects. However, due to experimenter error, 8 of 24 infants in the *different* condition had their test pairs dropped for this trial type. We conducted a repeated-measures ANOVA on the *Memory Check* test trial type, looking at the within-subjects variable of Relation (novel vs. familiar) and the between-subjects variables of Condition (habituation to *same* vs. *different*). Counter to our predictions, infants looked equally between these trials, an average of 19.14 s ($SD = 15.02$) at novel relations and 20.15 s ($SD = 13.97$) at familiar relations, and there was no significant effect of relation, $F(1, 38) < 1, p = .56$. A binomial comparison reflected this lack of discrimination, with 21 of the 40 infants with usable data for this trial type looking longer at the novel relation on this trial type, $p = .875$. There was no main effect of condition nor a relation by condition interaction.

8.1. Comparison to Experiment 1

The main question was whether infants would show differences in generalization to new objects between Experiment 1 and Experiment 2, indicating that the different types of labels had distinct effects. A repeated-measures ANOVA⁸ on the *New* test trial type, with the within-subjects variable of Relation (novel vs. familiar) and the between-subjects variables of Experiment (1 vs. 2), Condition (habituation to *same* vs. *different*), Test Order (novel relation or familiar relation first), and Sex (male vs. female), revealed that there was a main effect of relation where infants looked longer at the novel relations, $F(1, 96) = 9.90, p = .002, \eta_p^2 = .09$. There was no interaction of Relation and Experiment, $F(1, 96) = 1.01, p = .318$, nor an interaction of Relation and Condition, $F(1, 96) < 1, p = .938$. However, there was a three-way interaction of Relation $\times$ Experiment $\times$ Condition, where infants in Experiment 1 looked longer at the novel relation compared to the familiar one when they had habituated to *same*, but infants in Experiment 2 looked longer at the novel relation compared to the familiar one when they had habituated to *different*, $F(1, 96) = 9.32, p = .003, \eta_p^2 = .09$.

The next question was whether infants would show differences between the *Relational Label* trial in Experiment 1 and the *Object Label* trial in Experiment 2, indicating that the different types of labels had distinct effects. A repeated-measures ANOVA on the *Relational/Object Label* test trial type, with the within-subjects variable of Relation (novel vs. familiar) and the between-subjects

⁷ A pilot experiment implemented a similar procedure with 7- and 9-month-olds ($n = 32$) and, in contrast to Experiment 2, found null results on all trials: *New*, $F(1,29) < 1, p = .331$; *Object Label*, $F(1,30) = 2.58, p = .119$, and *Habituation*, $F(1,29) = 3.961, p = .056$. As in the current experiment, there was a Relation and Condition interaction for Habituation trials, where infants habituated to *same* performed better, $F(1,29) = 4.707, p = .039$.

⁸ To compensate for the unequal sample sizes in the ANOVA, JASP software uses weighted means and Type I sum of squares.

variables of Experiment (1 vs. 2), Condition (habituation to *same* vs. *different*), Test Order (novel relation or familiar relation first), and Sex (male vs. female), did not show a main effect of relation, $F(1, 93) < 1$, $p = .561$. There was no interaction of Relation and Experiment, $F(1, 93) < 1$, $p = .65$, nor any other main effects or interactions.

Finally, a repeated-measures ANOVA on the *Pair Habituation* test trial type, with the within-subjects variable of Relation (novel vs. familiar) and the between-subjects variables of Experiment (1 vs. 2), Condition (habituation to *same* vs. *different*), Test Order (novel relation or familiar relation first), and Sex (male vs. female) revealed a main effect of relation where infants looked longer at the novel relation, $F(1, 88) = 4.07$, $p = .047$, $\eta_p^2 = .04$. There was no interaction of Relation and Experiment, $F(1, 88) < 1$, $p = .396$. There was an interaction between Relation and Condition, such that, collapsing across experiments, infants habituated to *same* looked longer at the novel pair ($M = 18.28$) than at the familiar pair ($M = 14.48$), but infants looked equally between pairs in the *different* condition, ($M_{\text{Novel}} = 14.15$; $M_{\text{Familiar}} = 13.64$), $F(1, 88) = 5.40$, $p = .02$, $\eta_p^2 = .06$. There was a three-way interaction of Relation, Order and Sex, where male infants showed greater discrimination between relations when the novel relation was first in the test pair, and female infants showed the opposite pattern, $F(1, 88) = 5.65$, $p = .02$, $\eta_p^2 = .06$. There was also a between-subjects interaction of Experiment, Order and Sex, $F(1, 88) = 4.33$, $p = .04$, $\eta_p^2 = .05$: female infants had longer looking times in the familiar-first order in Experiment 1,

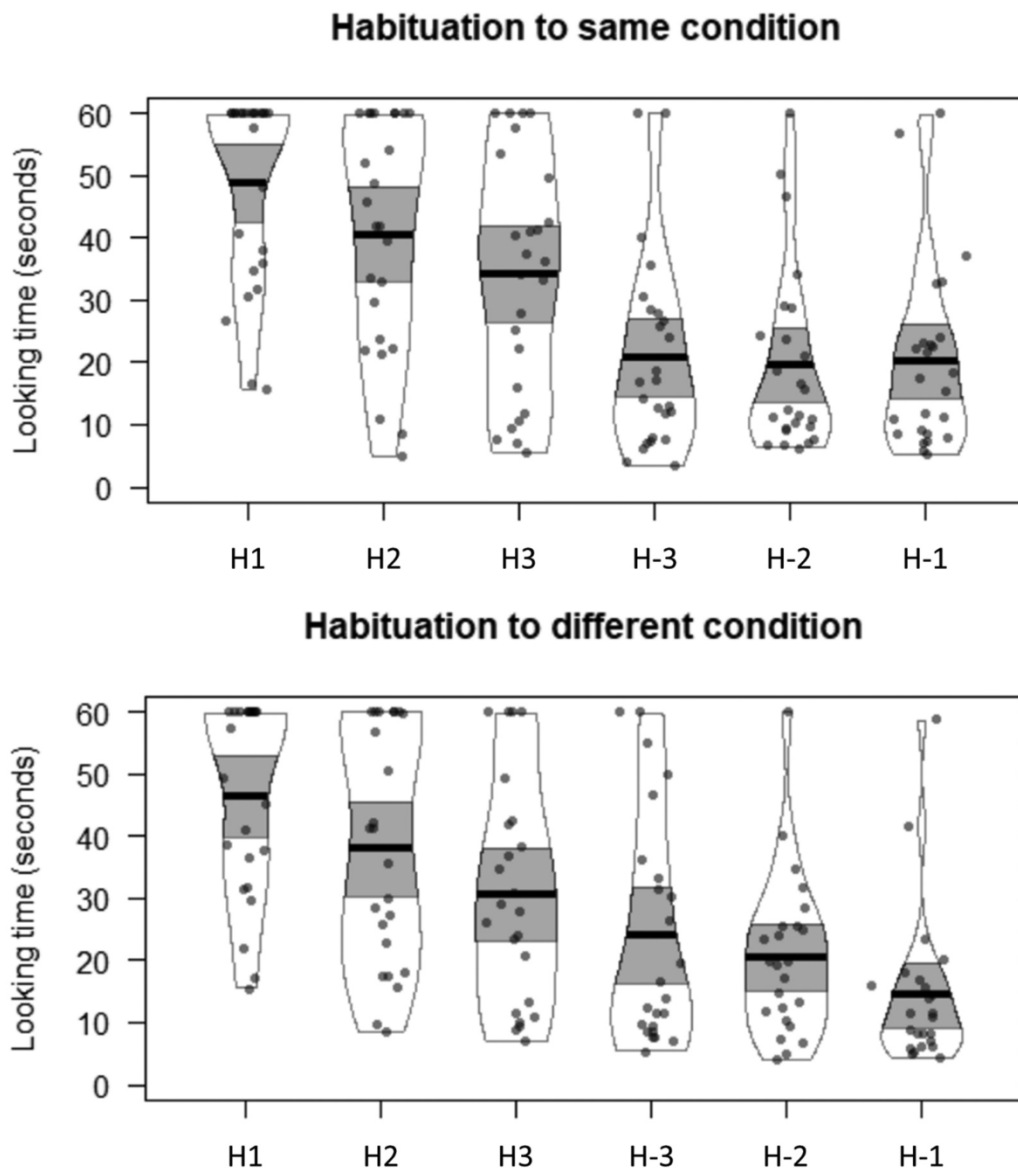


Fig. 6. Habituation trial looking times. Because the number of habituation trials was infant-controlled, ranging from 6 to 9 trials, the graphs represent the first three (H1–H3) and the last three trials (H-3, H-2, H-1). The thick central line in each box is the mean, and the upper and lower shaded portions represent the 95% Confidence Intervals (CIs) for this mean. Dots indicate the raw data points. The width of the bean indicates the density of the data distribution at a looking time value.

but the reverse was true in Experiment 2. There were no other main effects or interactions.

8.2. Comparison to [Ferry et al. \(2015\)](#)

We next asked whether labeling objects in Experiment 2 hindered generalization more than simply drawing attention to objects in [Ferry et al. \(2015\)](#). We conducted a repeated measures ANOVA on the *New* trials, using the within subjects factor of Relation and the between-subjects factors of Experiment and Condition, revealing a main effect of Relation, $F(1, 108) = 18.63, p < .001, \eta_p^2 = .147$. The effect of Relation did not depend on Experiment, $F(1, 108) < 1, p = .803$: looking times were greater for novel than for familiar trials in both studies, and the magnitude of the difference was comparable: in [Ferry et al. \(2015\)](#) ($M_{\text{Novel}} = 18.91$ s; $M_{\text{Fam}} = 13.09$ s) and in Experiment 2 ($M_{\text{Novel}} = 21.83$ s; $M_{\text{Fam}} = 16.44$ s). Relation did interact with Condition, $F(1, 108) = 5.029, p = .027$, showing that, across experiments, differences in looking time at the novel and familiar relation were more pronounced in the habituation to *different* condition ($M_{\text{Novel}} = 21.48$ s; $M_{\text{Fam}} = 13.03$ s) than in the *same* condition ($M_{\text{Novel}} = 19.14$ s; $M_{\text{Fam}} = 16.30$ s).

We next asked whether object labels disrupted relational processing as did object salience. We conducted a repeated measures ANOVA on the *Object Label/Salient Object* trials, examining the within-subject factor of Relation and the between-subject factors of Experiment ([Ferry et al., 2015](#) vs. Experiment 2) and Condition (habituation to *same* vs. *different*). This ANOVA showed no main effect of Relation, $F(1, 105) < 1, p = .837$. Relation and Experiment did not interact: infants in [Ferry et al. \(2015\)](#) looked near equally between relations ($M_{\text{Novel}} = 17.74$ s; $M_{\text{Fam}} = 17.23$ s) as did infants in the current experiment ($M_{\text{Novel}} = 17.88$ s; $M_{\text{Fam}} = 17.17$ s), $F(1, 105) < 1, p = .438$. There were no main effects or interactions. This suggests a local disruptive effect: in both studies, infants failed to discriminate the novel from the familiar relation on test trials that contained the objects that had been labeled (Experiment 2) or otherwise rendered salient ([Ferry et al., 2015](#)).

Finally, we examined the effect of object labels on Pair Habituation test trials, performing a repeated measures ANOVA with a within-subjects factor of Relation and the between-subjects factors of Experiment and Condition. As with *New* trials, there was a main effect of Relation, $F(1, 105) = 17.00, p < .001, \eta_p^2 = .139$. The effect of Relation did not depend on Experiment, $F(1, 108) = 1.91, p = .17$: looking times declined from novel to familiar trials in both [Ferry et al., \(2015\)](#) ($M_{\text{Novel}} = 18.91$ s; $M_{\text{Fam}} = 13.09$ s) and in Experiment 2 ($M_{\text{Novel}} = 14.45$ s; $M_{\text{Fam}} = 11.48$ s). There was a three-way interaction between Relation, Experiment and Condition, $F(1, 105) = 17.00, p < .001, \eta_p^2 = .139$, reflecting that infants in Experiment 2 showed clearer discrimination in the *same* condition, while infants in [Ferry et al. \(2015\)](#) discriminated in both conditions. There were no other effects or interactions.

8.3. Habituation patterns

Lastly, we examined looking time during the habituation phase. A repeated measures ANOVA examined looking time by the within-subjects factor of Habituation Trial (first three and last three) and the between-subjects factor of Condition (*same* and *different*). The ANOVA revealed a significant decline in looking across conditions, $F(5, 230) = 37.23, p < .001, \eta_p^2 = .45$ (see [Fig. 6](#)). On average, infants looked 48.20 s on their first habituation trial and 17.46 s on their last trial. There was no interaction between Habituation Trial and Condition, $F(5, 230) = 1.14, p = .339$, nor a main effect of Condition, $F(1, 46) < 1, p = .491$. Additional analyses show that Condition (habituation to *same* or *different*) did not predict total looking time during habituation, $F(1, 46) < 1, p = .394$, nor how many trials it took reach the habituation criterion, $X^2(3, N = 48) = 3.01, p = .39$, nor did it predict how many infants reached the habituation criterion, $X^2(1, N = 48) = 0.091, p = .763$. Additionally, infants' looking times on the first habituation trial was 50 s in the *same* condition and 46.4 s in the *different*, suggesting that the baseline looking time for *same* relations was not greater than for *different*, $t(46) = 0.84, p = .40$.

9. Discussion

In Experiment 2, infants failed to detect the relation in the *Object Label* trials, but did detect the relation in the *New* trials. This pattern is consistent with infant relational learning studies where the objects are not labeled but were rendered salient in other ways ([Ferry et al., 2015](#)). Therefore, we cannot conclude that the label itself was responsible for the increase in salience; the act of showing the object and speaking might have rendered it salient or not that speech was labeling the object.

The results revealed two other unexpected findings. First, there was an interaction between Relation and Condition for the *Pair Habituation* test type, which indicated that infants discriminated the novel and familiar relations in the *same* condition, but not in the *different* condition. This is surprising, in that none of the earlier infant studies in our lab found that *same* is more readily learned than *different*. However, other researchers have found such patterns ([Addyman & Mareschal, 2010](#); [Hochmann et al., 2018](#)). Another unexpected finding was that, in the *Memory Check* trials, infants completely failed to discriminate between pairs. It is puzzling that infants would succeed on the *Pair Habituation* trial but not the *Memory Check* trial, when the latter was designed to be an easier version of the former, with the novel pair also made of novel objects. It is unclear what drove this effect.

10. General discussion

These studies aimed to understand whether language would impact relational learning in preverbal infants as it does with older children, with relational labels improving performance and object labels hindering it. For calibration, we compared the results of labeling objects and relations with 12-month-olds to those of [Ferry et al. \(2015\)](#). In that study, 7- and 9-month-olds were first shown a set of (unlabeled) objects (to render them salient) and then received the same basic habituation experience as in the current

experiments. The results of Experiment 1 showed that hearing prior relational labels (*same* and *different*) did not improve relational generalization. Experiment 2 found that hearing prior object labels led to a parallel pattern as in [Ferry et al. \(2015\)](#). In both cases, infants were able to generalize to *New* pairs, but failed to show relational discrimination on test trials containing the labeled or salient objects.

Why did the infants fail to show any benefit from relational labels in Experiment 1? One possibility is that infants mistook the common label in Experiment 1 for an object noun instead of a relational label. Furthermore, labeling pairs as “these are *same*” instead of “these are the *same*” may have removed syntactic clue that signaled the *same* relation instead of a plural noun. This could have rendered the objects more salient and hinder infants’ detection of the relation.

Even if infants in Experiment 1 did not construe the terms “same” and “different” as object labels, they still may not have inferred the relational referent. This would be in line with evidence that relational concepts are slow to be acquired compared to object concepts. Infants’ early vocabulary learning skews towards concrete nouns ([Bornstein et al., 2004](#); [Gentner, 1982, 2005](#); [Gentner & Boroditsky, 2001, 2009](#); [Gleitman et al., 2005](#)). Although benefits of hearing highly frequent relational labels such as “in” and “on” have been found at 18 months ([Casasola, 2005](#)), it is not until three years that children show relational label benefits for *same* and *different* ([Christie & Gentner, 2014](#)).

Even in older children, relational language can be challenging. Although relational terms are common in English ([Asmuth & Gentner, 2005](#)), there is a protracted learning period for them, during which time children often make more concrete mappings—e.g., “taxi” as a yellow car rather than any car for hire ([Gentner, 2005](#); [Gentner & Rattermann, 1991](#); [Keil & Batterman, 1984](#)). Indeed, even when 4-year-old children do show an advantage of relational labels for relational abstraction, they often require additional support to do so. As an example, [Christie and Gentner \(2010\)](#) presented preschool age children with a pair of examples, both of which demonstrated the same relation, but differed in their concrete instantiations. Both pairs were given the same label, e.g., “This is a jiggy”. When the two examples were seen simultaneously, the children were able to derive the abstraction and correctly map the relation to another example. But when the examples were presented sequentially, children failed to learn the abstraction. Though both groups heard a common label for these relations, it was only the children in the simultaneous condition who successfully compared the two exemplars and derived the relational match. In light of the long learning period for relational language, it is perhaps not so surprising that 12-month-olds were not able to map a relational label to its referents.

11. Limitations and open questions

It is possible that the infants failed to learn the relations in Experiment 1 because three examples of *same* or *different* pairs—the number given to infants in the pretrial—were not enough. After all, the 3-year-olds in [Christie and Gentner’s \(2014\)](#) label-training study (described earlier) required a mean of 11 training trials (half *same* and half *different*) to reach criterion, and the 54% of 2 ½-year-olds who reached criterion required 19 trials (the rest failed even given 24 trials). Thus, although we cannot rule out the possibility that the 12-month-olds could learn *same* and *different*, the logistics of giving infants a sufficient number of trials prior to habituation are rather formidable.

A more promising future approach would be to add labeling during habituation to pretraining relations (with perhaps a few more examples than were used in the current Experiment 1). This would in effect combine [Christie and Gentner \(2014\)](#)’s two paths—pre-training labeled relations and using common labels to invite comparison.

Additionally, our design used words in the infants’ native language (e.g., “same”, “cup”) as our labels. It would be valuable to look at the role of labels with novel words that could play both roles, as novel labels have been shown to help with learning object categories (in infancy) and relations (in preschool) ([Althaus & Westermann, 2016](#); [Christie & Gentner, 2010, 2014](#); [Gentner & Namy, 1999](#); [Namy & Gentner, 2002](#); [Waxman & Braun, 2005](#)).

12. Conclusions

We found no evidence for a benefit of relational language on 12-month-olds’ ability to carry out relational abstraction, and only ambiguous evidence for a deleterious effect of object naming. Experiment 1 failed to show any effect of labeling the *same* and *different* relations. The results of Experiment 2 are consistent with the possibility that object labels interfered with relational processing of pairs containing these objects; but we cannot separate the labeling effect from a general salience effect. Overall, these findings are consistent with previous studies that have found relatively late acquisition of relational language. They are also consistent with studies that show that relational learning occurs prior to language acquisition ([Anderson et al., 2018](#); [Ferry et al., 2015](#)). It appears that interactions between language and relational processing – especially the beneficial effects of labeling relations – are slow to develop—possibly because relational language is itself slow to be learned. Ultimately, the results of our studies show that the interaction of relational learning and language is a protracted developmental process.

CRediT authorship contribution statement

Erin Anderson: Conceptualization, Investigation, Data curation, Formal analysis, Visualization, Writing – original draft. **Yin-Juei Chang:** Conceptualization, Investigation, Data curation, Software, Writing – review & editing. **Susan Hespos:** Conceptualization, Investigation, Writing – review & editing, Funding acquisition, Supervision. **Dedre Gentner:** Conceptualization, Writing – review & editing, Funding acquisition, Supervision.

Acknowledgments

This research was supported by the National Science Foundation grants [BCS-1423917, BCS-1729720] awarded to Susan J. Hespos and Dedre Gentner, by NSF grant [SBE-0541957] awarded to the Spatial Intelligence and Learning Center (SILC), by Office of Naval Research grant [N00014-16-1-2613] awarded to Dedre Gentner, and by a US Department of Education, Institute of Education Sciences Training Grant (Multidisciplinary Program in Education Science) [R305B140042] awarded to Erin Anderson. We are indebted to parents who agreed to have their infants participate in the research. We thank members of the Infant Cognition Lab for their help, especially Shelley Powers and Apoorva Shivaram. All data from the experiments reported in this paper will be archived in the [Supplementary materials](#).

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at [doi:10.1016/j.infbeh.2021.101666](https://doi.org/10.1016/j.infbeh.2021.101666).

References

- Addyman, C., & Mareschal, D. (2010). The perceptual origins of the abstract same/different concept in human infants. *Animal Cognition*, 13(6), 817–833. <https://doi.org/10.1007/s10071-010-0330-0>
- Althaus, N., & Westermann, G. (2016). Labels constructively shape object categories in 10-month-old infants. Interrelations between non-linguistic and linguistic representations of cognition and action in development. *Journal of Experimental Child Psychology*, 151, 5–17. <https://doi.org/10.1016/j.jecp.2015.11.013>
- Anderson, E. M., Chang, Y.-J., Hespos, S., & Gentner, D. (2018). Comparison within pairs promotes analogical abstraction in three-month-olds. *Cognition*, 176, 74–86. <https://doi.org/10.1016/j.cognition.2018.03.008>
- Asmuth, J.A., & Gentner, D. (2005). *Context sensitivity of relational nouns*. (<http://www.escholarship.org/uc/item/8p9181j1>).
- Bergelson, E., & Swingle, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences of the United States of America*, 109(9), 3253–3258.
- Bergelson, E., & Swingle, D. (2013). The acquisition of abstract words by young infants. *Cognition*, 127(3), 391–397.
- Bornstein, M. H., Cote, L. R., Maital, S., Painter, K., Park, S.-Y., Pascual, L., ... Vyt, A. (2004). Cross-linguistic analysis of vocabulary in young children: Spanish, Dutch, French, Hebrew, Italian, Korean, and American English. *Child Development*, 75(4), 1115–1139. <https://doi.org/10.1111/j.1467-8624.2004.00729.x>
- Casasola, M. (2005). Can language do the driving? The effect of linguistic input on infants' categorization of support spatial relations. *Developmental Psychology*, 41(1), 183–192. <https://doi.org/10.1037/0012-1649.41.1.183>
- Chang, Y.-J., Wang, S.-W., & Way, B. (2018). *BLT: experimental program for online coding babies' looking time*. (github.com/NUinfantlab/BLTProgram).
- Christie, S., & Gentner, D. (2010). Where hypotheses come from: Learning new relations by structural alignment. *Journal of Cognition and Development*, 11(3), 356–373.
- Christie, S., & Gentner, D. (2014). Language helps children succeed on a classic analogy task. *Cognitive Science*, 38(2), 383–397.
- Csibra, G., Hernik, M., Mascaro, O., Tatone, D., & Lengyel, M. (2016). Statistical treatment of looking-time data. *Developmental Psychology*, 52(4), 521–536.
- Ferry, A. L., Hespos, S. J., & Gentner, D. (2015). Prelinguistic relational concepts: Investigating analogical processing in infants. *Child Development*, 86(5), 1386–1405.
- Fisher, C. (1996). Structural limits on verb mapping: The role of analogy in children's interpretations of sentences. *Cognitive Psychology*, 31(1), 41–81. <https://doi.org/10.1006/cogp.1996.0012>
- Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2016). Wordbank: An open repository for developmental vocabulary data. *Journal of Child Language*, 1–18.
- Gentner, D. (2003). Why we're so smart. In *Language in mind: Advances in the study of language and thought* (pp. 195–235). MIT Press.
- Gentner, D. (2005). The development of relational category knowledge. In L. Gershkoff-Stowe, & D. H. Rakison (Eds.), *Building Object Categories in Developmental Time: Building Object Categories in Developmental Time* (pp. 245–275). Hillsdale, NJ: Erlbaum.
- Gentner, D. (2010). Bootstrapping the mind: Analogical processes and symbol systems. *Cognitive Science*, 34(5), 752–775.
- Gentner, D., Anggoro, F. K., & Klibanoff, R. S. (2011). Structure-mapping and relational language support children's learning of relational categories. *Child Development*, 82(4), 1173–1188.
- Gentner, D., & Boroditsky, L. (2001). Individuation, relativity and early word learning. In M. Bowerman, & S. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 215–256). Cambridge University Press.
- Gentner, D., & Boroditsky, L. (2009). Early acquisition of nouns and verbs: Evidence from Navajo. In V. Gathercole (Ed.), *Routes to language: Studies in honor of Melissa Bowerman* (pp. 5–36). New York, NY: Taylor & Francis.
- Gentner, D., & Christie, S. (2008). Relational language supports relational cognition in humans and apes: A response to Penn, Holyoak & Povinelli. *Behavioral and Brain Sciences*, 31, 137–183.
- Gentner, D., & Christie, S. (2010). Mutual bootstrapping between language and analogical processing. *Language and Cognition*, 2(2), 261–283.
- Gentner, D., & Medina, J. (1998). Similarity and the development of rules. *Cognition*, 65, 263–297.
- Gentner, D., & Namy, L. L. (1999). Comparison in the development of categories. *Cognitive Development*, 14(4), 487–513.
- Gentner, D., & Rattermann, M. J. (1991). Language and the career of similarity. In *Perspectives on language and thought: Interrelations in development* (pp. 225–277). Cambridge University Press.
- Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. A. Kuczaj (Ed.), *Language development: Vol. 2. Language, thought and culture* (pp. 301–334). Hillsdale, NJ: Erlbaum.
- Gentner, D., Shao, R., Simms, N., & Hespos, S. (2021). Learning same and different relations: Cross-species comparisons. *Current Opinion in Behavioral Sciences*, 37, 84–89. <https://doi.org/10.1016/j.cobeha.2020.11.013>
- Gentner, D., & Toupin, C. (1986). Systematicity and surface similarity in the development of analogy. *Cognitive Science*, 10(3), 277–300.
- Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, 15, 1–28.
- Gleitman, L. R. (1990). The structural sources of verb meanings. *Language Acquisition*, 1(1), 3–55.
- Gleitman, L. R., Cassidy, K., Nappa, R., Papafragou, A., & Trueswell, J. C. (2005). Hard words. *Language Learning and Development*, 1(1), 23–64. <https://doi.org/10.1207/s15473341lild0101.4>
- Hespos, S. J., Ferry, A. L., & Rips, L. J. (2009). Five-month-old infants have different expectations for solids and liquids. *Psychological Science*, 20(5), 603–611. <https://doi.org/10.1111/j.1467-9280.2009.02331.x>
- Hespos, S. J., Grossman, S. R., & Saylor, M. M. (2010). Infants' ability to parse continuous actions: Further evidence. *Neural Networks: The Official Journal of the International Neural Network Society*, 23(8–9), 1026–1032.
- Hochmann, J.-R., Carey, S., & Mehler, J. (2018). Infants learn a rule predicated on the relation same but fail to simultaneously learn a rule predicated on the relation different. *Cognition*, 177, 49–57.

- Hofstadter, D., & Sander, E. (2013). *Surfaces and essences: Analogy as the fuel and fire of thinking*. New York, NY: Basic Books.
- Hoyos, C., Shao, R., & Gentner, D. (2016). The paradox of relational development: Could language learning be (temporarily) harmful? In D. Grodner, D. Mirman, A. Papafragou, J. Trueswell, J. Novick, S. Arunachalam, S. Christie, & C. Norris (Eds.), *Proceedings of the 38th annual conference of the cognitive science society*. Cognitive Science Society.
- Keil, F. C., & Batterman, N. (1984). A characteristic-to-defining shift in the development of word meaning. *Journal of Verbal Learning & Verbal Behavior*, 23(2), 221–236. [https://doi.org/10.1016/S0022-5371\(84\)90148-8](https://doi.org/10.1016/S0022-5371(84)90148-8)
- Loewenstein, J., & Gentner, D. (2005). Relational language and the development of relational mapping. *Cognitive Psychology*, 50(4), 315–353.
- Markman, A. B., & Gentner, D. (1993). Structural alignment during similarity comparisons. *Cognitive Psychology*, 25, 431–467.
- Namy, L. L., & Gentner, D. (2002). Making a silk purse out of two sow's ears: Young children's use of comparison in category learning. *Journal of Experimental Psychology: General*, 131(1), 5–15.
- Paik, J. H., & Mix, K. S. (2006). Preschoolers' use of surface similarity in object comparisons: Taking context into account. *Journal of Experimental Child Psychology*, 95(3), 194–214.
- Penn, D. C., Holyoak, K. J., & Povinelli, D. J. (2008). Darwin's mistake: Explaining the discontinuity between human and nonhuman minds. *The Behavioral and Brain Sciences*, 31(2), 109–178. <https://doi.org/10.1017/S0140525x08003543>
- Rattermann, M. J., & Gentner, D. (1998). The effect of language on similarity: The use of relational labels improves young children's performance in a mapping task. In *Advances in analogy research: Integration of theory & data from the cognitive, computational, and neural sciences* (pp. 274–282). New Bulgarian University.
- Son, J., Smith, L. B., Goldstone, R., & Leslie, M. (2012). The importance of being interpreted: Grounded words and children's relational reasoning. *Frontiers in Developmental Psychology*, 3, 1–12 (PMCID: PMC3296077).
- Tyrrell, D. J., Stauffer, L. B., & Snowman, L. G. (1991). Perception of abstract identity/difference relationships by infants. *Infant Behavior and Development*, 14, 125–129. [https://doi.org/10.1016/0163-6383\(91\)90059-2](https://doi.org/10.1016/0163-6383(91)90059-2)
- Waxman, S. R., & Braun, I. (2005). Consistent (but not variable) names as invitations to form object categories: New evidence from 12-month-old infants. *Cognition*, 95(3), B59–B68. <https://doi.org/10.1016/j.cognition.2004.09.003>